THE PREDICTION OF RUNNING SPEEDS ON URBAN ARTERIAL STREETS

NOVEMBER 1971 - NUMBER 20

BY

K. A. SHACKMAN

JHRP

JOINT HIGHWAY RESEARCH PROJECT

PURDUE UNIVERSITY AND

INDIANA STATE HIGHWAY COMMISSION

THE PREDICTION OF RUNNING SPEEDS ON URBAN ARTERIAL STREETS

NOVEMBER 1971 - NUMBER 20

BY

K. A. SHACKMAN

Final Report

THE PREDICTION OF RUNNING SPEEDS ON URBAN ARTERIAL STREETS

TO:	J. F. McLaughlin, Director Joint Highway Research Project	November 23, 1971
		Project: C-36-17II
FROM:	H. L. Michael, Associate Director Joint Highway Research Project	File: 8-4-35

The attached Final Report titled "The Prediction of Running Speeds on Urban Arterial Streets" has been authored by Kenneth A. Shackman, Graduate Assistant in Research on our staff. The research has been supervised by Professor H. L. Michael.

The research reported herein concerns the development of a prediction model for running speeds on urban arterial streets outside of the Central Business District. The determination of the major causes of stop delay and their average values (in time) are also included.

The report is presented as fulfillment of the objectives of the Plan of Study titled as this Final Report and approved by the Advisory Board on September 15, 1970. It is presented for the record and for acceptance.

Respectfully submitted,

Harold L. Michael

Harold L. Michael
Associate Director

HLM:ms

cc:	W. L. Dolch	R. H. Harrell	W. T. Spencer
	W. H. Goetz	M. L. Hayes	J. A. Spooner
	W. L. Grecco	R. D. Miles	N. W. Steinkamp
	M. J. Gutzwiller	J. W. Miller	H. R. J. Walsh
	G. K. Hallock	C. F. Scholer	K. B. Woods
	M. E. Harr	M. B. Scott	E. J. Yoder

Final Report

THE PREDICTION OF RUNNING SPEEDS ON URBAN ARTERIAL STREETS

by

Kenneth A. Shackman
Graduate Assistant in Research

Joint Highway Research Project

Project: C-36-17II

File: 8-4-35

Conducted By

Joint Highway Research Project
Engineering Experiment Station
Purdue University

In Cooperation With

Indiana State Highway Commission

Purdue University
Lafayette, Indiana
November 23, 1971

ACKNOWLEDGMENTS

The author wishes to express his gratitude to the Directors of the Joint Highway Research Project for having enabled him to continue his education.

To Mr. Edward J. Kannel for his help, when nothing seemed to be going as planned, and for his moral support and confidence the remainder of the time, the author is indebted.

The esteem of the author goes to Professor H. L. Michael, Head of Transportation and Urban Engineering, School of Civil Engineering, for his guidance in the selection of this topic, for his constructive comments throughout the study, and for his critical review of the manuscript; to Professor V. L. Anderson, Department of Mathematics and Statistics for his advice and review of the statistical analyses and review of the manuscript; and to Professor K. W. Heathington for his advice and review of the manuscript.

Acknowledgment is also given to Mr. George Stafford for his help during the data collection phase of the project.

Thanks go to the many Traffic Engineers, City Engineers, Chiefs of Police, and Police Traffic Officers who assisted in the data collection.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES. vii

ABSTRACT viii

INTRODUCTION 1

Past Research in This Area. 2

PURPOSE. 5

DATA COLLECTION. 7

Determination of Sample Size for Speed Measurements. 7
Selection of Study Sections 9
Data Collection Procedures. 21

DATA ANALYSIS. 28

Original Variables. 28
Preliminary Testing 32
Stepwise Regression 34
Phase 1. 34
Phase 2a 35
Phase 2b 35
Phase 3a 37
Phase 4a 37
The Prediction Model. 37

ANALYSIS OF STOP DELAY 48

CONCLUSION 63

BIBLIOGRAPHY 69

APPENDICES

Appendix A: Discussion of the Foster-Burr Q-Test. 71

TABLE OF CONTENTS (Continued)

Page

Appendix B: Normality Test for Section Speeds . 75
Appendix C: Normality Test for Residuals. . . . 78

LIST OF TABLES

Table		Page
1	Population Groupings of Cities in U.S.A. and Indiana	10
2	Study Cities and Their Classification	12
3	Selected Test Sections With Pertinent Data	16
4	Variables Tested in Development of the Model	29
5	Summary of Homogeneity Test Results	33
6	Stepwise Regression Summary Table	39
7	Results of Testing the Model	44
8	Average Delay Time Per Section (Secs)	50
9	Average Delay Time Per Section (Secs)	55
10	Average Delays Due to Traffic Signals, Railroad Crossings, and Other Factors (Secs)	59
11	Summary of Speed Influencing Factors	67

Appendix Table		
B1	Normality Test for Section Speeds (Off-Peak and Peak)	76
C1	Normality Test for Residuals	79

LIST OF FIGURES

Figure		Page
1	Frequency of Section Length	14
2	Inventory Questionnaire	22
3	Data Sheet for Travel Time Study.	24
4	Summary Sheet for Travel Time Study	25
5	Model Inputs.	36
Appendix Figure		
B1	Frequency of Section Speeds	75
C1	Frequency of Residuals.	78

ABSTRACT

Shackman, Kenneth Alan. M.S.C.E., Purdue University, January, 1972. The Prediction of Running Speeds on Urban Arterial Streets. Major Professor: Harold L. Michael.

The purpose of this research project was the development of a model for predicting running speeds on urban arterial streets. To supplement the model the major causes of stop delay and their average values were also evaluated. Using these results one can estimate the change in average running speed and stop delay time between an existing facility and a proposed or improved facility. The estimate of the benefits gained, in terms of time, speed, or money, could then be contrasted with the proposed cost.

The data collection was conducted in nine Indiana cities having populations between thirty and one hundred and fifty thousand; the average car method was the procedure employed for collecting the speed-delay data. Pneumatic tube traffic counters were installed to determine the volumes using the streets being studied.

The prediction model was formulated using stepwise regression. It utilizes eleven roadway and traffic condition variables combined to form ten terms; it has an R^2 of 0.5926 and a standard error of estimate of 3.417.

The major influences of speed are:

1. The length of the section
2. The width of the approach
3. The number of signalized intersections
4. The number of unsignalized intersections
5. The number of alleys
6. The population of the city
7. The number of railroad crossings
8. Whether or not parking was allowed on the driver's right
9. The volume
10. The lowest G/C ratio of the traffic signals included in the section
11. Whether or not good progression of the traffic signals existed

The average values of stop delay were calculated from data collected in the speed-delay studies conducted in nine cities. Traffic signals were found to be the major source of stop delays, contributing over eighty percent of the total.

INTRODUCTION

In a recent publication (5) it was stated "that at least as many drivers selected the least 'effort' routes as the least time routes to a shopping center." Just what makes least "effort" routes so inviting? Basically, three elements of high quality of service are considered: 1) safety, 2) constant (or gradually changing) direction, and 3) a constant, reasonable speed. This particular study was concerned with the latter element, speed.

Though a constant speed is desirable the probability of it occurring on an urban arterial street is low, the presence of traffic control devices and a high volume of traffic preventing it. This is true even on streets equipped with progressive signal systems. With proper traffic engineering controls and an adequate design, however, a reasonable average running speed for this type of facility can be obtained.

Average running speed data are widely utilized to determine benefits gained by improvement of the quality of traffic service. This has been accomplished by performing speed-delay studies both before and after implementation of improvements and contrasting the running speeds. Examples of areas where this has been done include measuring

"the overall effectiveness of traffic control devices", in general (10), and traffic signal timing revisions, in particular, and modifications concerning parking prohibitions, development of one-way street networks, etc. In presenting such accrued benefits of savings in time (or its monetary value) and a usual concurrent reduction of accidents, the engineer has shown the taxpayer, as well as the politician, the value of traffic engineering improvements.

The goal of this study has been to determine a means of predicting the average speed on sections of urban arterials without performing speed-delay studies. This was accomplished by developing a model for predicting the average running speed, employing experimentation and analysis similar to Hejal (7). By estimating average speeds, the time and manpower requirements are reduced, as well as the cost. Indirectly, the major causes of speed reduction and their relative influence have also been determined.

Past Research In This Area

During the past decade there has been some research in the travel time or speed and delay areas. Basically, three different approaches have been tried. They are:

1. Computer Simulation
2. Before and After Studies
3. Models - Theoretical and Empirical

Computer simulation appears to be promising; however, this approach has not been perfected to the stage where the results are both workable and representative of reality. There are computer solutions for parts of the problem, such as predicting speeds of vehicles passing through signalized intersections (1) and determining delaying effects of curb parking maneuvers on traffic (19). Though these are steps in the needed direction, the latter, as well as others, did not produce results representative of reality. It is probable that the development of a realistic simulation program that will encompass all aspects of traffic friction is not in the immediate future. The development of an empirical procedure, however, appears to be a quickly attainable goal since the formulation period would be short and the results might be easily applicable and reasonably accurate.

Edwards and Kelcey (3) reviewed the many parameters of frictional effects on travel time. There were many variables involved in their work in downtown areas. The procedure that was followed was to do "before and after studies" for determining the quantitative effect of individual improvements. These were statistically tested for significance and a Network Assignment Model was applied to determine the overall efficiency of the network. The method worked quite effectively, but it has the disadvantage

of being very expensive. Such a method, however, used outside the Central Business District (CBD) where fewer variables would have to be measured and analyzed might be economically reasonable.

Another researcher (5) included change of speed in route evaluation by correlating "cost, comfort, and safety" with a number. The number, however, also was based on change of direction and change of time.

$$N_t = \Delta \text{ time} \times \Delta \text{ speed} \times \Delta \text{ direction}$$

The resulting number does not mean anything by itself; it must be compared with another to show relative improvement along a particular section. Unfortunately, this method lumps all the influential factors of quality of service together. The analyzer has no way of determining the cause of quality deterioration, whether it is due to traffic volume or the facility's design characteristics.

Finally, in Hejal's report (7) a regression analysis was used to obtain a relationship between six traffic parameters and the running speed for a two-lane rural highway. Data were collected using the average car method, noting the causes of interference and including other factors or variables from a road inventory. The results obtained compared well with reality and at the same time had a high coefficient of multiple determination (R^2). The method proved to be both economical and simple to perform.

PURPOSE

The specific purpose of this research was to develop a model for predicting the average running speeds of vehicles on arterial streets in fringe areas of the CBD. This, it was assumed, could be achieved by investigating the influence of traffic volume, traffic controls, and design features of the street. This model has two functions: one, to predict the speed that an average vehicle will travel along a facility; and two, to evaluate selected causes of speed reduction.

It is hoped that the proposed model will be useful on all arterial streets, both existing and proposed. On existing streets it should emphasize to the engineer where improvements should be made and their relative benefit. On proposed streets it should indicate during the design stage the average running speed of the vehicles that will use it under anticipated conditions and, if this speed proves to be too low, where design changes need to be made. Another possible use will be the determination of appropriate speed limits.

The engineer today can estimate a desired quality of service from an evaluation of the capacity of a facility. The results of this research will be another tool,

hopefully one which is easier and more accurate of application. The following statement taken from the "Highway Capacity Manual-1965" (9) is a warrant for conduct of this study: "Insufficient data are available to attempt to develop correction factors (on capacity and speed) for such mid-block influences as curb parking." It was apparent that more research was needed on factors affecting average speed.

DATA COLLECTION

Determination of Sample Size for Speed Measurements

At the outset of the data collection stage of the study a section sample size had to be chosen. This sample would consist of a number of running speed observations along one section of roadway in one direction. The decision was made to calculate the size using a statistical approach (more observations would be an inefficient use of both time and money).

Before the calculation of the sample size could be completed certain assumptions had to be made. First, it was assumed that the running speeds recorded on the test sections would be normally distributed. If this were so then the following formula could be used for the calculation:

$$n = z_{\frac{\alpha}{2}}^{2} \ \sigma^{2} / E^{2} \qquad (14)$$

where

n = sample size
z = table value based on the chosen confidence level
σ = standard deviation of the population
E = acceptable error

The value of α, the probability of a Type I error chosen was 0.05 and the value of E chosen was $\pm$ 2.0 mph. The

value used for the standard deviation of the population, σ, had to be determined from a pilot study since there were no recent data available.

The pilot study was conducted in West Lafayette, Indiana on Northwestern Avenue (U.S. 52) between Stadium Avenue and Lindberg Road, a distance of approximately 1.1 miles. Data were collected for two directions on three separate Tuesdays in three months in 1970. The license plate method of data collection was used during the early afternoon of each day. A fifteen minute period was allocated to each direction for collection, providing between thirteen and twenty-five pieces of usable data for analysis.

The analysis showed that the standard deviation ranged from 1.669 to 4.851 (requiring sample sizes from three to twenty-two). However, the majority of the values were centered around 3.2, producing a sample size of ten. It was at this point that a homogeneity of variance test was conducted to determine whether the six values of σ^2 were equal. A Statistical Program, DATASUM which includes the Foster-Burr Test (Q-Test) was used to determine this (See Appendix A for description of this test). The data passed the Foster-Burr Test. Hence, it was decided to accept the hypothesis that the variances (and standard deviations) were equal. A sample size of ten was then chosen for use in data collection.

Selection of Study Sections

Once the sample size for each section had been chosen (ten observations) it was then necessary to determine the number of cities and the number of sections in each city to be examined.

At the outset of this research project it was decided that the work should be limited to cities whose population ranged from thirty to one hundred and fifty thousand people. There were two reasons for this decision: economy and expedience. Economy refers to conserving the limited resources of both time and money. This was achieved by restricting the data collection to nine Indiana cities in order to limit the resources expended in traveling between cities. Cities with populations less than thirty thousand were not included because the characteristics of their arterial streets are considerably different than such streets in larger cities, i.e. the fringe area is very small or non-existent, homogeneous street sections are very short, and traffic engineering is often not used. The ground for not examining cities over one hundred and fifty thousand population is that they are too few in number (Table 1). According to Table 1 eighty-five percent of all Indiana cities with populations greater than thirty thousand fall within the specified range. At the national level the figures are just as impressive: eighty-three percent of all cities over twenty-five thousand have

TABLE 1

POPULATION GROUPINGS OF CITIES IN U.S.A. AND INDIANA

U.S.A. (1960 Census)		Indiana (1970 Census)	
Number of cities with populations > 10000	1899	Number of cities with populations > 30000	20
Number of cities with populations > 25000	765	Number of cities with populations between 30000 and 150000	17
Number of cities with populations between 25000 and 250000	714		
Number of cities with populations between 25000 and 100000	633		

populations between twenty-five and one hundred thousand, while ninety-three percent of all cities over twenty-five thousand have populations between twenty-five and two hundred and fifty thousand.

In order to economize limited resources and to aid some of the smaller cities of the state by supplying them with study data, it was decided to collect all data within the state of Indiana. After organizing the seventeen cities that were included in the specified range into the three classes it was found that the second and third classes each contained three cities. The decision was then made to use three cities from the first class, thereby creating nine cities to be investigated (Table 2). In Class I the preference of one city over another included consideration of its proximity to Purdue.

Each section investigated was observed under two traffic conditions. The first was the afternoon off-peak and the second was the evening peak period. The evening peak was chosen instead of the morning because it tends to be better defined and greater in magnitude. However, using these two conditions covers most of the spectrum of problem traffic conditions except the possible situations where a two-way street carries high volumes in both directions during the noon lunch hour.

The judgment was made to research ten sections in each city to provide a sizeable number of observations at the

TABLE 2

STUDY CITIES AND THEIR CLASSIFICATION

Class	Population Range	City
I	30000 - 49999	Lafayette Kokomo Marion
II	50000 - 99999	Anderson Muncie Terre Haute
III	100000 - 149999	Hammond South Bend Evansville

completion of data collection.

The choice of streets actually studied was accomplished by visual inspection. In making this judgment three factors were employed. The first was location within the city. This was restricted to the fringe area of the Central Business District. The second criterion was section length. In chosing the test sections an attempt was made to maintain a minimum length of one half mile; however, this proved quite difficult at times due to the third selection factor, homogeneity. Each street used had to be uniform along the length analyzed in the following four aspects:

1. Street Type (i.e. one-way or two-way)
2. Parking
3. Approach Width
4. The Absence of STOP and YIELD Signs in the Section

For the most part, satisfying the above conditions, with the exception of length, was not exceedingly difficult. Of the ninety sections investigated sixty-three percent were one-half mile or longer while slightly over eighty-three percent were 0.45 miles or longer (Figure 1). In the cases where the section length was less than four-tenths of a mile (only seven of ninety sections) the number of major causes of speed reduction, e.g. traffic signals and railroad crossings, was kept to a minimum (usually one) in order to guarantee that the test vehicle

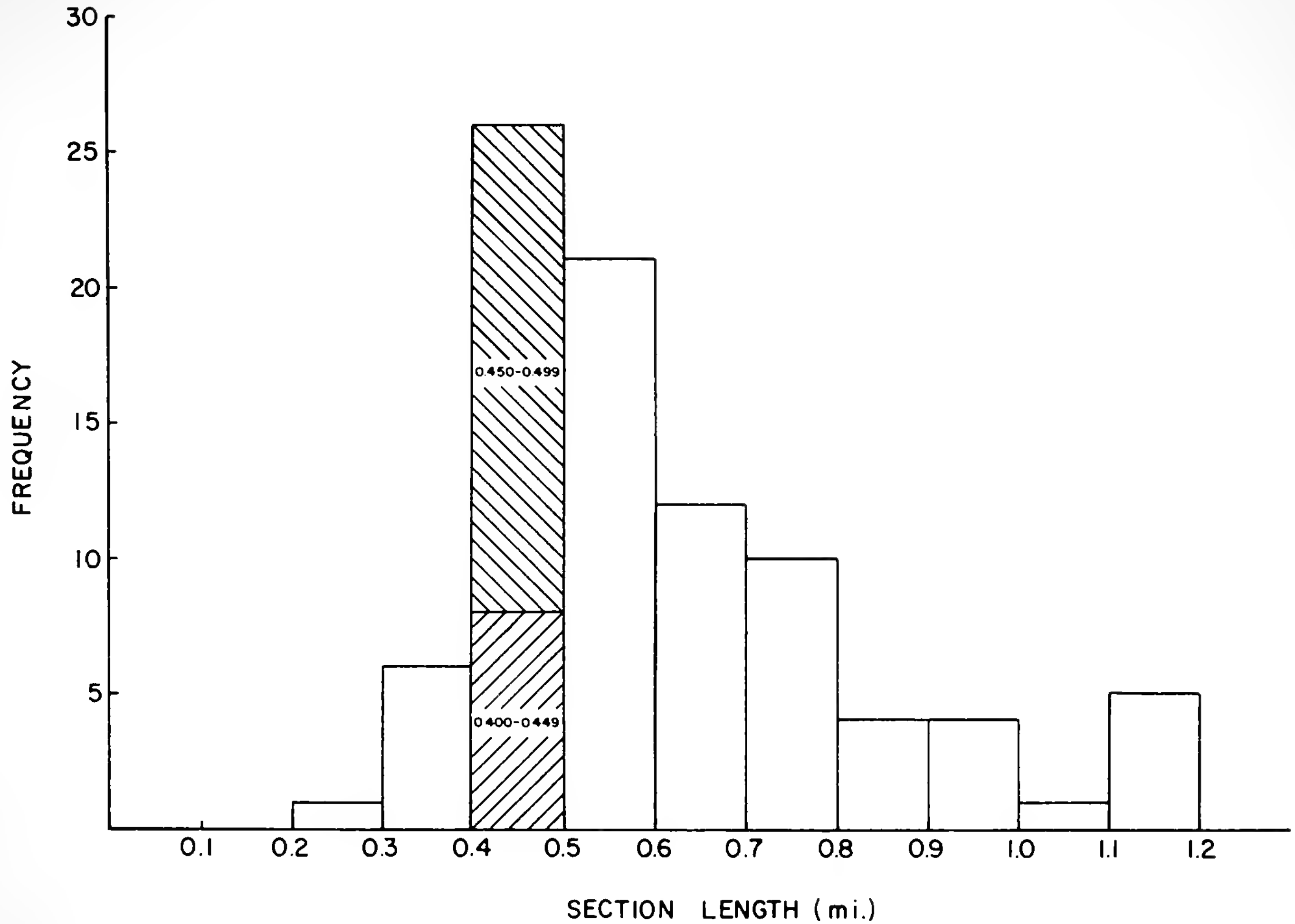

FIGURE 1. FREQUENCY OF SECTION LENGTH

could cruise and not merely accelerate and decelerate. Refer to Table 3 for a listing of the streets selected.

Although ten sections of arterial streets were desired in each of the nine cities, only eight adequate sections could be found in Marion. As a consequence an additional two sections were used in Kokomo, a similar sized city in the same area of the state.

Although the posted speed limits of the sections ranged from 20 mph to 40 mph, less than eight percent of the streets had speeds different than 30 mph. The average speeds on those streets with a posted speed less than 30 mph were found to be about thirty miles per hour, while those streets with a posted speed greater than 30 mph were approximately thirty-five miles per hour.

During the off-peak traffic condition observations were taken on all ninety sections; however, during the peak traffic condition observations were taken on only sixty-two sections. This was a result of the absence of a peak volume on some streets during the evening peak period. In some cases the section was a one-way street of a pair, while in others it was one direction of a two-way street. As a consequence, of the one hundred and eighty conditions that could possibly be investigated only one hundred and fifty-two actually were analyzed.

TABLE 3

SELECTED TEST SECTIONS WITH PERTINENT DATA

City	Street	Sec. No.	Terminal Points		Length (Mi.)	Peak 15 Min. Vol.	Surrounding Area Description
Lafayette	Union	7	14th	1st Barrel	0.411	309	Res,Com
	Main	9	South	Earl	1.187	158	Res,Com
	Main	11	Earl	South	1.199	110	Res,Com
	Northwestern	35	Lindberg	Stadium	1.097	94	Res
	Northwestern	36	Stadium	Lindberg	1.106	304	Res
	Teal	57	9th	18th	0.486	186	Res,Com
	Teal	58	18th	9th	0.486	196	Res,Com
	Columbia	59	10th	Main	0.298	179	Res
	9th St.	64	Potomac	Virginia	0.477	93	Res
	9th St.	65	Virginia	Potomac	0.484	214	Res
Kokomo	Apperson Way	13	North	Jefferson	0.483	60	Res
	Apperson Way	16	Jefferson	North	0.489	97	Res
	Sycamore	17	Phillips	Webster	0.518	63	Res
	Sycamore	18	Webster	Phillips	0.520	126	Res
	Jefferson	29	Wabash	Korby	0.415	76	Res,Com Ind
	Jefferson	30	Korby	Wabash	0.416	60	Res,Com Ind
	Phillips	72	Elm	Walnut	0.555	35	Res,Com Ind
	Phillips	73	Walnut	Elm	0.556	113	Res,Com Ind
	Markland	74	Plate	Buckeye	0.885	176	Com
	Markland	75	Buckeye	Plate	0.887	156	Com
	Morgan	89	Apperson Way	Webster	0.476	165	Res,Com
	Morgan	90	Webster	Apperson Way	0.476	193	Res,Com

TABLE 3 (Continued)

City	Street	Sec. No.	Terminal Points		Length (Mi.)	Peak 15 Min. Vol.	Surrounding Area Description
Anderson	Broadway	19	Grand	Cross	1.185	249	Res,Com
	Broadway	20	Cross	Grand	1.180	237	Res,Com
	Cross	21	Silver	Lowell	0.441	117	Res
	Cross	22	Lowell	Silver	0.444	197	Res
	8th St.	23	Henry	Brown	0.745	181	Res
	8th St.	24	Brown	Henry	0.752	135	Res
	Madison	25	5th	Van Buskirk	0.634	246	Res
	Madison	26	Van Buskirk	5th	0.633	121	Res
	Meridian	27	17th	25th	0.538	209	Res,Com Ind
	Main	28	24th	16th	0.564	193	Res,Com Ind
Muncie	Wheeling	31	Cowing	Highland	0.681	229	Res,Com
	Wheeling	32	Highland	Cowing	0.683	294	Res,Com
	McGalliard	33	Lanewood	Walnut	0.718	171	Res,Com Ind
	McGalliard	34	Walnut	Lanewood	0.717	195	Res,Com Ind
	Tillotson	45	Burnell	Euclid	0.364	134	Res
	Tillotson	46	Euclid	Burnell	0.359	174	Res
	Memorial	47	Elliot	Utica	0.646	98	Res,Com
	Memorial	48	Utica	Elliot	0.646	160	Res,Com
	Jackson	49	Madison	Ohio	0.481	180	Res,Com
	Main	50	Wolf	Monroe	0.506	149	Res,Com

TABLE 3 (Continued)

City	Street	Sec. No.	Terminal Points		Length (Mi.)	Peak 15 Min. Vol.	Surrounding Area Description
Terre Haute	Poplar	37	22nd	12th	0.843	104	Res,Com Ind
	Poplar	38	12th	22nd	0.838	185	Res,Com Ind
	13th St.	39	Maple	6th	0.625	102	Res,Ind
	13th St.	40	6th	Maple	0.625	132	Res,Ind
	25th St.	41	8th	Maple	0.498	84	Res
	25th St.	42	Maple	8th	0.496	91	Res
	Maple	43	25th	15th	0.741	57	Com,Ind
	Maple	44	15th	25th	0.747	62	Com,Ind
	Ohio	62	13th	18th	0.428	195	Res
	Walnut	63	18th	13th	0.426	107	Res
Marion	Wabash	51	Kem	Quarry	0.516	120	Res
	Wabash	52	Quarry	Kem	0.516	147	Res
	Washington	53	McKinley	Grant	0.550	138	Res,Com
	Washington	54	Grant	McKinley	0.547	158	Res,Com
	Washington	55	11th	26th	0.937	176	Res,Com
	Adams	56	26th	11th	0.936	184	Res,Com
	W. 2nd St.	60	Baldwin	Geneva	0.389	142	Res,Com
	W. 2nd St.	61	Geneva	Baldwin	0.397	115	Res,Com

TABLE 3 (Continued

City	Street	Sec. No.	Terminal Points		Length (Mi.)	Peak 15 Min. Vol.	Surround-ing Area Descrip-tion
South Bend	Sample	66	Camden	Edison	0.758	219	Com,Ind
	Sample	67	Edison	Camden	0.757	172	Com,Ind
	Edison	68	Beutter	Pyle	0.535	160	Res
	Edison	69	Pyle	Beutter	0.534	143	Res
	Portage	70	Bergan	Queen	0.636	92	Res
	Portage	71	Queen	Bergan	0.636	105	Res
	Jefferson	80	Jacob	Ironwood	0.646	228	Res.
	Jefferson	81	Ironwood	Jacob	0.644	145	Res
	Washington	82	Laurel	Birdsell	0.476	126	Res,Com
	Washington	83	Birdsell	Laurel	0.473	91	Res,Com
Hammond	Columbia	14	Kenwood	Drackert	0.547	194	Res,Com
	Columbia	15	Drackert	Kenwood	0.551	354	Res,Com
	165th St.	76	Columbia	Indianapolis	0.996	187	Com,Ind
	165th St.	77	Indianapolis	Columbia	0.997	235	Com,Ind
	Kennedy	78	173rd	169th	0.492	184	Com
	Kennedy	79	169th	173rd	0.489	322	Com
	Summer	84	Hammond Valve	Berkley	0.729	161	Res,Ind
	Summer	85	Berkley	Hammond Valve	0.713	142	Res,Ind
	Michigan	86	Driveway 2416	Indianapolis	0.561	150	Ind
	Michigan	87	Indianapolis	Driveway 2416	0.553	189	Ind

TABLE 3 (Continued)

City	Street	Sec. No.	Terminal Points		Length (Mi.)	Peak 15 Min. Vol.	Surround-ing Area Descrip-tion
Evansville	Franklin	1	Oakley	Main	0.511	140	Res,Com
	Franklin	2	Main	Oakley	0.501	103	Res,Com
	Garvin	3	Franklin	Louisiana	0.571	124	Res,Com
	Governor	4	Missouri	Illinois	0.571	283	Res,Com
	Washington	5	Kentucky	Lodge	0.491	241	Res
	Washington	6	Lodge	Kentucky	0.481	134	Res
	Columbia	8	3rd	6th	0.326	181	Res,Com
	Columbia	10	7th	Wabash	0.446	140	Res
	Columbia	12	Wabash	7th	0.451	114	Res
	Columbia	88	6th	3rd	0.321	157	Res,Com

Data Collection Procedures

The data that were collected for this project may be divided into three general categories: 1) speed-delay data, 2) volume data, and 3) section inventory data. A section refers to either a one-way street or one direction of a two-way street.

The initial step of research in a city was the obtainment of a city street map and the labeling of the perimeter of the Central Business District (CBD) and the arterial streets. This was usually done by contacting the City Traffic Engineer; however, in some cases it was necessary to speak to either the Police Traffic Officer or the City Engineer. This also proved to be a good opportunity to speak to either the Chief of Police or the Police Traffic Officer and inform them of the type of study that was to be done.

Once the preliminaries were complete the actual data collection began. This required driving to the CBD perimeter and checking each arterial for length and homogeneity. The former was done using a vehicle equipped with a calibrated Stewart-Warner Survey Speedometer (Odometer) while the latter was accomplished by answering the pertinent questions of the Inventory Questionnaire (Figure 2). The remainder of the questionnaire was completed at the researcher's convenience.

Inventory Questionnaire

Check the appropriate spaces

City	________________
Street Name	________________
Type of Street	One-Way ______ Two-Way ______
Parking	No Parking ________
	Parking (one side) ______
	N E S W side of street
	Parking (both sides) ______
Number of Railroad Crossings	________________
Number of Driveways	Commercial ____ Private ____
Section Direction	North ________ East ________
Approach Width	________________
Lane Markings	Yes ________ No ________
Section Direction	South ________ West ________
Approach Width	________________
Lane Markings	Yes ________ No ________
Number of Signalized Intersections	________________
Number of Intersections Other than Signalized	________________
Number of Alleys	________________
G/C Ratios	

Intersecting Street	G/C Ratio
________________	________________
________________	________________
________________	________________
________________	________________
________________	________________

FIGURE 2. INVENTORY QUESTIONNAIRE

After a street was chosen, Streeter Amet Trafficounters (Models RC or RCT) were installed in each direction. These pneumatic tube traffic volume counters were placed so that the volume that passed that point was representative of what travelled on the whole section.

Upon completion of this phase a speed-delay study was begun, using the average car method. In brief, this method involved one of the researchers driving the test vehicle along the section at what he considered the average speed of the traffic stream while simultaneously measuring, using a stop watch, the total time of the trip. At the same time, a recorder timed all stops and noted both these and other forms of delay. The data sheet completed by the recorder is shown in Figure 3. Upon returning to the office the researcher completed the data sheet and proceeded to complete the Summary Sheet (Figure 4).

This procedure was followed on each run with ten runs performed under two traffic conditions on each section, during the off-peak period and again during the evening peak period. To determine the time and magnitude of the peak period a traffic volume counter was left in position for twenty-four consecutive hours between Monday and Friday. However, this method had to be modified on many sections due to equipment malfunction or damage to the pneumatic tubing by street cleaning equipment. The modification applied was the installation of the traffic

TRAVEL TIME STUDY

TRIP NO. ____________

ROUTE ________________________ DIRECTION _____________________

TRIP STARTED AT _____________ HRS _________________________
LOCATION MILEAGE

TRIP ENDED AT _______________ HRS _________________________
LOCATION MILEAGE

REFERENCE		STOPS			SLOWS	
POINT	TIME	LOCATION	CAUSE	DELAY SEC.	LOCATION	CAUSE

TOTAL TRIP TIME ___________ TOTAL TRIP LENGTH ____ OVERALL SPEED __

RUNNING TIME _____________ STOPPED TIME __________ RUNNING SPEED __

DATE _____________________ WEATHER __________________________________

FIGURE 3. DATA SHEET FOR TRAVEL TIME STUDY

Summary Sheet

Route ______________________ Length ______________________

Between ____________________ And ______________________

Date _______ Time ______ AM/PM to _____ AM/PM Weather ___________

Major Cause for Delay	Trip No.	No. Times Stopped	Total Sec. Stopped	No. Times Slowed	Average Speed
TOTALS					
AVERAGES PER TRIP					

Date ______________________ Compiled By ______________________

FIGURE 4. SUMMARY SHEET FOR TRAVEL TIME STUDY

volume counter during the late morning hours and the removal of it the same day during the early evening hours.

If this type of study is to be a success the people being observed must not know what is taking place. The test vehicle is important in this matter. The vehicle used for the data collection was a 1962 Chevrolet four door sedan; in no way was the vehicle conspicuous.

The Inventory Questionnaire (Figure 2) was the only original form used in the study, the others being standard forms from the Manual of Traffic Engineering Studies (10). It functioned as a substitute for a condition diagram of the section using both qualitative and quantitative questions. All the required measurements could be completed by one person with the exception of that of the approach width. Usually three width measurements were taken, one at each end and one in between, using a metallic tape. The G/C ratios were usually supplied by the City Traffic Engineer; however, in cases where they were not, the measurements were made and the percentages rounded to the nearest integral number.

An inherent problem in the average car method arose in data collection. During off-peak measurements there were periods when the test vehicle was not included within visible range of other vehicles. The driver of the test vehicle then had to make the decision of choosing a speed that he felt was average. To wait outside the section for

other vehicles usually does not alleviate the problem - all too often under low volume conditions a group of other vehicles does not materialize. It was sometimes difficult under such conditions for the driver of the test car to be certain he drove at the speed of the average vehicle.

The bulk of the data collection was done between February and May of 1971 and only during dry weather conditions. It should also be noted that the data was collected on weekdays, with the exception of one Saturday afternoon.

DATA ANALYSIS

Original Variables

Table 4 is a listing of the variables used in the development of the model. Of the twenty-two variables investigated, only eleven are actually included in the proposed model. It should also be noted that one of the twenty-two variables, Driver, was added just prior to the choosing of the model. This was done to try to improve the model, but which it did to only a very small degree.

Information for many of the variables was obtained directly from the Inventory Questionnaire. Data for the twenty first variable, Progression, however, resulted from value judgments on the part of the researcher. Value judgments present difficulties because other researchers might have made different decisions. In the case of a section where there were zero or one signalized intersection, the section was considered to have good progression. Where the section had more than one signalized intersection, the researcher on the basis of speed progression actually obtained during the runs made the judgment as to whether or not good progression was present.

The twenty-second variable, Driver, represents whether Driver A or Driver B was at the wheel of the test vehicle.

TABLE 4

VARIABLES TESTED IN DEVELOPMENT OF THE MODEL

Variable Number	Variable Name	Description and Information Coded
1	Running Speed	The length of the section divided by the time in motion for one trip along that section (mph).
2	Peak	Denotes whether measurements were done during the P.M. Peak period or not. Off-Peak = 0 P.M.Peak = 1
3(A)	Section Length	The length of the section in miles.
4	Street Type	Denotes whether the section is a one- or two-way street. One-way = 0 Two-way - 1
5(B)	Approach Width	The curb to curb width (face of curb) in feet for one-way streets and the curb to centerline width in feet for two-way streets.
6	Lane Marking	Used for multilane facilities only. No = 0 Yes = 1
7(C)	Signalized Inter.	The number of signalized intersections included in the section.
8(D)	Unsignalized Inter.	The number of unsignalized intersections included in the section.
9(E)	Alleys	The number of alleys included in the section.
10	Private Driveways	The number of private driveways included in the section.
11	Commercial Driveways	The number of commercial driveways included in the section.

TABLE 4 (Continued)

Variable Number	Variable Name	Description and Information Coded
12	Day	The day of the week that the data was collected. Sun = 1 Mon = 2 : Sat = 7
13(F)	Pop.	Denoted the population class 30000-49999 = 0 50000-99999 = 1 100000-149999 = 2
14(G)	Railroad Crossings	The number of individual crossings (not the number of tracks) included in the section.
15	Parking	Variables 15, 16, and 17 are Parking dummy variables whose relationship is:
16	Parking	15 16 17 0 0 0 Parking on both sides of the section
17(H)	Parking	1 0 0 No parking on either side of the section 0 1 0 Parking on the left side of the section (one-way streets) 0 0 1 Parking on right side of the section.
18 (I)	Volume	The fifteen minute volume which was concurrently measured with the observation (one trip).
19(J)	Lowest G/C ratio	The lowest G/C ratio of all the signalized intersections included in the section.
20	Average G/C ratio	The average of all the G/C ratios included in the section.

TABLE 4 (Continued)

Variable Number	Variable Name	Description and Information Coded
21(K)	Good Progression	A value judgment of whether good progression existed in those sections which contained two or more signalized intersections. No = 0 Yes = 1 (also used for sections with zero or one signalized intersection)
22	Driver	Denotes driver of the test vehicle Driver A = 0 Driver B = 1

Range of Variable Measurements

Variable Number	Range
3	0.298 miles - 1.199 miles
7	0 traffic signals - 6 traffic signals
13	30000 people - 150000 people
14	0 crossings - 2 crossings
18	24 vehicles - 354 vehicles
19	25% - 100%
20	25% - 100%

Upon examination of the covariance matrix it was found that one driver drove, on the average, nine-tenths of a mile per hour faster than the other.

A listing of the limitations of the individual variables, where this is relevant, is given at the end of Table 4.

Preliminary Testing

As data collection progressed it was felt necessary to determine if the data indicated a homogeneous population. To accomplish this the Q-Test contained in Statistical Program DATASUM was again employed. Table 5 is a summary of these results, including the results of the tests made at the completion of the data collection stage. The tests made during data collection indicated homogeneity while those made upon completion of data collection indicated heterogeneity. A possible reason for the failures for the data at completion might be the non-normality (see Appendix B) and not the heterogeneity of the data (4).

At this stage of the analysis it was realized that interaction between the variables was quite high and an attempt was made to determine it, using two-way ANOVA tests. Unfortunately, due to the incompleteness of the cells, it was found impossible to perform these tests. Therefore, regression analysis was employed, using cross products to investigate the interaction.

TABLE 5

SUMMARY OF HOMOGENEITY TEST RESULTS

Number of Variances Tested	Q-Test (0.05 Level)
(Tests Made During Data Collection)	
6	Passed
12	Passed
16	Passed
20	Passed
24	Passed
44	Passed
(Tests Made Upon Completion of Data Collection)	
62 (Peak)	Failed
90 (Off-Peak)	Failed
152 (Total)	Failed

Stepwise Regression

Towards the completion of the data collection phase, investigation of Statistical Program BMD 2R (Stepwise Regression) (17) began. Using three hundred and sixty observations, different F-ratios and Tolerances were examined for deviation between models created. Due to a lack of deviation in the models the following values were chosen and utilized throughout the analysis: F-ratio = 1.75 and Tolerance = 0.001.

Phase 1

Using a sample of six hundred observations and containing only the first seventeen variables (see Table 4), the input for a model was created. Nine of these variables were chosen for use in forming transgenerated variables, e.g. squares and cross products. These nine variables were:

1. Section Length
2. Approach Width
3. Signalized Intersections
4. Unsignalized Intersections
5. Alleys
6. Railroad Crossings

7.-9. Parking Dummy Variables

Those transgenerated variables which were either removed from the model during its development or never entered because of low F-ratios were eliminated from further consideration.

Soon afterwards another two hundred and ten observations were added and the process continued. The logarithms of some of the original variables were also examined during this phase of the regression analysis. Later another three hundred observations were added to the model's input.

Phase 2a

Phase 2a was an attempt at utilizing the most significant variables found from analyses of the eight hundred and ten observations and the eleven hundred and ten observations of the previous phase. At the same time variables concerning volume and traffic signal phasing were added. The input for the Phase 2a models is diagrammatically shown in Figure 5. At the completion of Phase 2a the R^2 value was 0.6679 for a model containing twenty-nine variables (original and transgenerated).

Phase 2b

Simultaneous to the running of the Phase 2a models Phase 2b was attempted. The Phase 2b input is shown in Figure 5. The function of this phase was to check whether or not the three hundred additional observations of the eleven hundred and ten observation model had caused certain variables to become more significant. At the termination of Phase 2b an R^2 value of 0.6919 had been reached using thirty-six variables. Because the final Phase 2b model

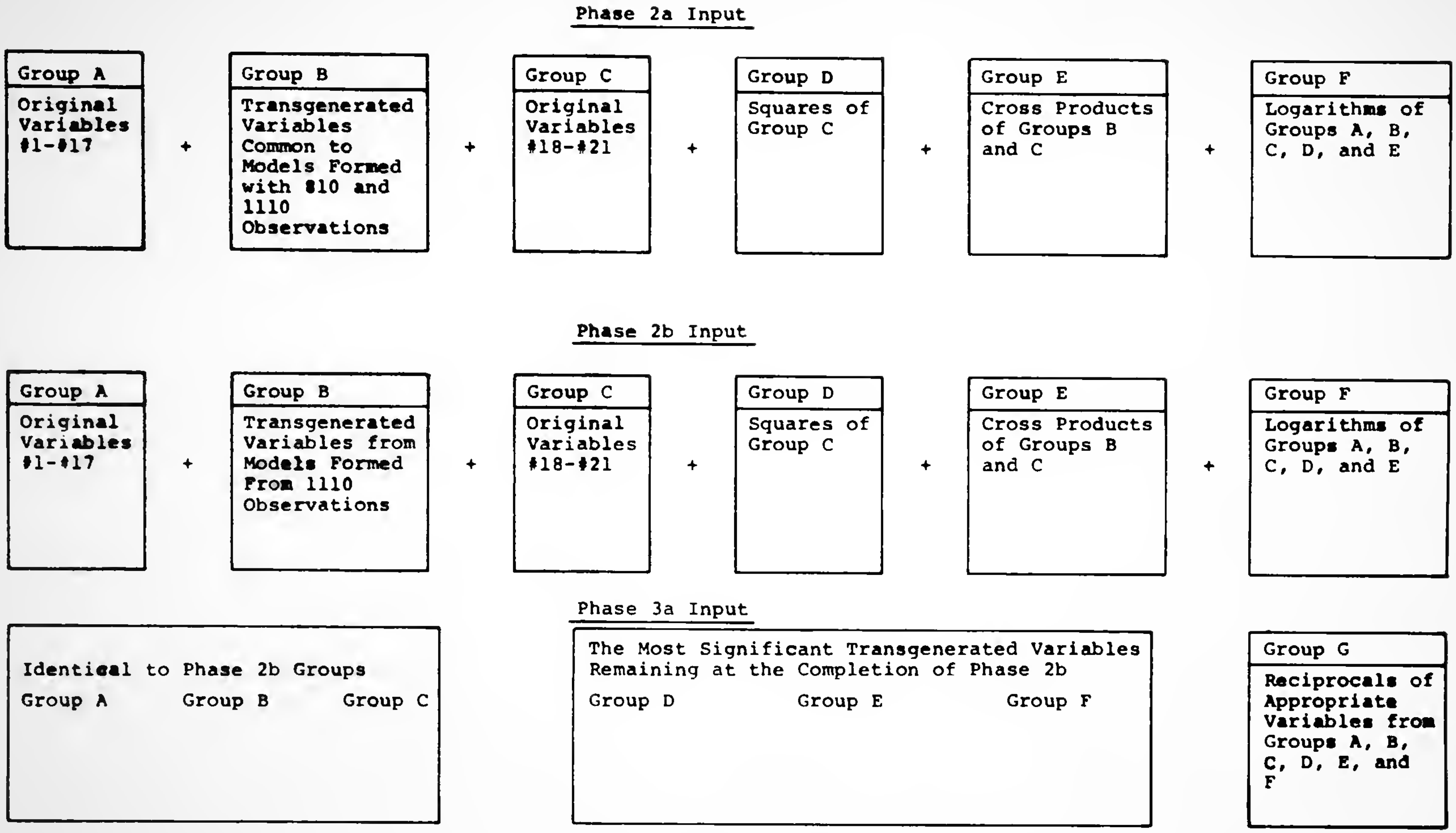

FIGURE 5. MODEL INPUTS

appeared more promising than its Phase 2a counterpart, it was decided to enter Phase 3a using the most significant variables from the Phase 2b model.

Phase 3a

Phase 3a was based on the same input as Phase 2b except for the addition of the more significant transgenerated variables entered during Phase 2b and the reciprocals of some of the variables. These reciprocals were also used to form new cross products. Although many new transgenerated variables were examined the value of R^2 remained relatively unchanged.

Phase 4a

Phase 4a was the stage in which illogical transgenerated variables were removed from the model. In addition to the twenty-one original variables, another ten logical (logical in individual effect) variables were selected from the most significant transgenerated variables entered into each of five Phase 3a models. It was from this phase that the final prediction model was developed.

The Prediction Model

The model that was chosen contains ten terms and utilizes eleven original variables. It represents the Phase 4a model where each term increased the value of R^2 by approximately one percent. The model is shown below in

test format, followed by the Summary Table from the stepwise regression program (Table 6).

$$\hat{Y} = 19.62500 + 10.38018X_3 + 0.07562X_5 - 0.70139X_7$$
$$- 0.22178X_8 + 1.25800X_{13} - 0.03118X_{18} + 7.38855X_{19}$$
$$- 0.18590X_{22} - 0.24199X_{23} + 0.01815X_{24}$$

where

$\hat{Y}$ = predicted average running speed

$$X_{22} = \frac{X_8 X_{14}}{X_3}$$

$$X_{23} = \frac{X_9 X_{17}}{X_3}$$

$$X_{24} = X_{18} X_{21}$$

See Table 4 for definitions of all other variables.

In Chapter 3 of Draper and Smith (2), "The Examination of Residuals," it states:

> We can see that ... the residuals e_i are the differences between what is actually observed, and what is predicted by the regression equation - that is, the amount which the regression equation has not been able to explain. Thus we can think of the e_i as the observed errors if the model is correct. Now in performing the regression analysis we have made certain assumptions about the errors; the usual assumptions are that the errors are independent, have zero mean, a constant variance, σ^2, and follow a normal distribution.

Appendix C contains the frequency plot of the residuals and the Kolmogorov-Smirnov Test for normality. The frequency plot looked normal and the test accepted the

TABLE 6

STEPWISE REGRESSION SUMMARY TABLE

Step No.	Variable Entered	Multiple R	R^2	Increase In R^2	F Value to Enter Or Remove	Standard Error of Estimate
1	19	0.5159	0.2661	0.2661	550.5366	4.5772
2	22	0.5951	0.3541	0.0880	206.5654	4.2909
3	3	0.6504	0.4230	0.0689	180.9292	4.0570
4	21	0.6896	0.4755	0.0526	151.7938	3.8692
5	23	0.7095	0.5034	0.0279	85.0990	3.7660
6	13	0.7297	0.5324	0.0290	93.7886	3.6557
7	18	0.7458	0.5562	0.0238	80.9227	3.5628
8	24	0.7532	0.5673	0.0112	39.0363	3.5188
9	-21*	0.7532	0.5673	-0.0001	0.2747	3.5180
10	7	0.7594	0.5767	0.0095	33.7673	3.4805
11	5	0.7635	0.5829	0.0062	22.2784	3.4562
12	8	0.7698	0.5926	0.0097	35.8412	3.4170

* Note: A minus sign means that the variable was removed.

hypothesis at a significance level of 0.05. The mean of the residuals was zero. To test the independence of the residuals two different approaches were used, neither producing favorable results. Each approach used a different ordering of the "runs" test outlined in Draper and Smith (2). The first ordering was a time approach. This was done by chronologically ordering the residuals of ten randomly chosen workdays. The results produced, however, were unsatisfactory. The ordering of the predictions (from lowest to highest) and their corresponding residuals was the second approach. Again, the results were the same.

A study of the plot of the residuals and the observations (running speeds) showed that as the speeds approached the limits of the observation range, the residuals tended to be of one sign. Therefore, if a chosen speed were at either extremity all the residuals would be of one sign, independent of time ordering. Because all speeds taken during a single workday were on only a few sections, the chosen runs on a single workday (the basis of this ordering approach) tended to have similar speeds, sometimes at one of the extremities.

In many instances, residuals are ordered with respect to one of the independent variables. This was not possible in this study because the variables were discrete over a narrow range. In place of one of the independent variables,

the prediction $\hat{Y}$ had to be used. This approach did not prove independence because the ordering from lowest to highest provided large groups of residuals of similar sign.

Though the tests to determine whether or not the residuals were independent failed, there may be good reasons for this result. The test used is most sensitive to independence for models of excellent fit. The model tested explains about sixty percent of the variation of the total sums of the squares of the running speed. Finally, as already noted, the ordering choices for which the residuals were tested presented difficulties.

The multiple correlation coefficient (R^2) for the model is 0.5926. There are three basic reasons why this value is not higher. The first reason is that the number of terms appearing in the model was kept to a minimum for calculation ease. Adding another ten terms would have raised the value of R^2 by only another six percent and some of these were not individually logical. It should be noted that the highest value of R^2 achieved during the analysis was 0.7022 using a model containing forty-two terms.

The second and probably the most important reason is the variability of drivers in the traffic stream. Many of the factors that influence a driver's actions are known but cannot be measured either quantitatively or qualitatively. To give some examples: 1) motivation

factors, such as attitude; 2) emotional factors, such as anger and fear; and 3) other modifying factors, such as fatigue, climate, and time of day. Motorists who are fatigued and inattentive drive differently than they normally would and some may even try to compensate for these factors. Unfortunately, in making this type of study the researcher has great difficulty determining which of these factors are influencing (or to what extent) the drivers in the traffic stream which he is studying.

The final reason is variability due to the method of data collection - the average car method. Late in the analysis stage the question of variability between the two test drivers arose. The twenty second variable, Driver, was then created to examine this. First, it was tested in a stepwise regression program with the original variables and proved to be more important than some of them. This new variable was tested with the terms of the final model. The result of this test was a negligible improvement in the explanation (R^2); therefore, it was not included. Another reason for this method's variability is the already discussed difficult driver selection of an average speed under low volume conditions.

After the selection of the model the decision was made to test it using part of the data collected. Ten of the sections were chosen from the data by randomly ordering peak and off-peak conditions, then randomly

ordering the section numbers and matching the first ten of each. The model was tested with coefficients to five places and the coefficients rounded to three places. Since the results were adequate for both coefficient conditions (Table 7), it was decided to use the model with rounded coefficients to simplify calculation.

To determine the accuracy of the model as a predictor, the standard deviation of each prediction, $S_{\hat{Y}}$, was estimated. The value of $S_{\hat{Y}}$ could be calculated using an equation similar to the equation for the one independent variable case, shown below:

$$s_{\hat{Y}}^2 = s_E^2 \left[1 + \frac{1}{n} + \frac{(X-\bar{X})^2}{\Sigma x^2} \right]$$

where

X = the independent variable

$x = X-\bar{X}$

The assumption was made that the third term in brackets was small and it was known that 1/n (1/1520) was very small, so that

$$s_{\hat{Y}}^2 \approx s_E^2$$

The standard error of the estimate (3.417) then approximates the actual value of $s_{\hat{Y}}$.

TABLE 7

RESULTS OF TESTING THE MODEL

Sect. No.	Peak Period	Street Name	City	Average Speed of Observations (mph)	Predictions of Model Utilizing Five Place Coefficients (mph)	Predictions of Model Utilizing Three Place Coefficients (mph)
89	No	Morgan	Kokomo	22.792	22.310	22.313
31	Yes	Wheeling	Muncie	24.508	27.132	27.141
22	No	Cross	Anderson	29.534	32.044	32.054
41	No	25th St.	Terre Haute	28.022	28.871	28.880
37	No	Poplar	Terre Haute	24.239	28.563	28.570
51	No	Wabash	Marion	30.316	27.770	27.776
42	Yes	25th St.	Terre Haute	26.341	29.281	29.291
80	No	Jefferson	South Bend	30.532	29.562	29.577
68	Yes	Edison	South Bend	35.128	33.363	33.374
20	Yes	Broadway	Anderson	31.720	31.921	31.936

TABLE 7 (Continued)

Sect. No.	Peak Period	Street Name	City	Average Speed of Observations (mph)	Predictions of Model Utilizing Five Place Coefficients (mph)	Predictions of Model Utilizing Three Place Coefficients (mph)
					Number of Predictions Whose Deviation Were Included in the ± 2 mph Range	
					5	5
					Average Deviation From The Average of the Section Observations	
					1.921	1.923
					Maximum Deviation from the Average of the Section Observations	
					4.324	4.332

The model in its final form using alphabetical subscripts is shown below.

$$\hat{Y} = 19.625 + 10.380X_A + 0.076X_B - 0.701X_C - 0.222X_D + 1.258X_F - 0.031X_I + 7.389X_J - 0.186\frac{X_D X_G}{X_A} - 0.242\frac{X_E X_H}{X_A} + 0.018X_I X_K$$

where

$X_A = X_3$

$X_B = X_5$

$X_C = X_7$

$X_D = X_8$

$X_E = X_9$

$X_F = X_{13}$

$X_G = X_{14}$

$X_H = X_{17}$

$X_I = X_{18}$

$X_J = X_{19}$

$X_K = X_{21}$

Though eleven variables were entered into the model, eleven others were not. It might be of interest to examine some of them now. Proceeding in the order of significance, the next variable that would have been

entered was Street Type (one-way or two-way) the coefficient of which was illogical. Surprisingly enough, the variable which denoted whether or not the prediction was being made for the peak condition entered into the model very late in the analysis, indicating a lack of significance, and had a small coefficient. Another variable which originally was hypothesized as being influential was Commercial Driveways. This variable also entered the model late in the regression analysis.

ANALYSIS OF STOP DELAY

The prediction of average running speed is a reasonable means of determining and evaluating some of the causes of speed reduction; however, to justify where and what improvements should be made it is usually necessary to have knowledge concerning any stop delay time. This section will be devoted to pointing out those locations where stop delays occurred in this study and their average values. Equipping the traffic engineer with both a running speed prediction model and average stop delay data will enable him to estimate where improvements should be made and the total benefits accrued therefrom.

A major part of this study was devoted to the measurement of stop delay, i.e. the time elapsed while the test vehicle was at a complete stop. On the sections investigated in this research project these stop delays were caused by three different factors:

1. Traffic Signals
2. Railroad Crossings
3. Miscellaneous Delays

Utilizing the data from this study, it was found that Traffic Signals had the greatest influence on stop delay

for both the off-peak and peak condition. For the off-peak condition the percentage of the total stop delay attributed to each factor was as follows: 81.90%, Traffic Signals; 15.50%, Railroad Crossings; and 2.60%, Miscellaneous Delays. During the peak period the same factors had values of 84.91%, 12.47%, and 2.62%, respectively.

Miscellaneous Delays contains any other causes of stop delay except by traffic signals and railroad crossings; however, during this study it was produced by left turns, right turns, school crossings, stopped vehicles in the roadway, and pedestrians crossing the roadway. The major and most common delay producing element of this factor was the left turn (36.80% of this factor's stop delay during the off-peak period and 74.90% during the peak period).

Tables 8 and 9 contain the delay data collected during the study. Table 8 is a tabulation of the off-peak data while Table 9 is a tabulation of the peak data. The delays recorded are the averages of the stop delays encountered during the ten observations along each section for each condition. Within Miscellaneous Delays the delays have been presented as the average delay per mile as opposed to the localized effects of Traffic Signals and Railroad Crossings.

A listing of the important statistics developed from the stop delay data is shown in Table 10.

TABLE 8

AVERAGE DELAY TIME PER SECTION (SECS)

Off-Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
7		16.30						0		0
9				24.52						0
11				34.15						0
35		6.59								1.09
36		5.43								2.29
57		23.23								0
58		0.70								0
59	0							23.97		0
64	0									0
65	0									0
13				4.80				0		0
16				19.09				0		0
17	0									0
18		8.30								0
29			11.60					0		0
30		10.68						0		0
72				18.43				0		0

TABLE 8 (Continued)

Off-Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
73			13.73					0		0
74						12.63			33.60	2.92
75							20.99		0	0.65
89		8.01						0.42		0
90			15.29					0		0
19			3.88					0		0
20			7.84					0		0
21	0							14.81		0
22	0							0		0
23			16.48					0		1.88
24			15.29					0		0
25		4.79								0
26		7.12								0
27			16.40						0	0
28				9.99					0	3.01
31			11.03							0
32			10.80							0
33		6.62						1.65		2.55
34	0							10.75		1.35

TABLE 8 (Continued)

Off-Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
45	0									0
46	0									0
47			12.93							1.11
48			15.70							0
49		3.17								0
50		9.72								0
37				5.38						0
38				13.47						0
39		5.98						43.88		0
40			24.58					0		0.32
41		7.39								0
42		0.98								0
43	0								10.22	0
44		13.20							2.55	0
62		0								0
63			13.59							0
51		5.90								0
52		6.45								1.47
53		8.62								0

TABLE 8 (Continued)

Off-Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
54		3.73								0
55		0						0		0
56	0							0		2.21
60	0									0
61		17.74								0.76
66	0									0
67	0									0
68	0									0
69	0									0
70	0									0
71	0									0
80				2.45						0.54
81			0							0
82		5.72								2.42
83		4.44								10.07
14			11.12							0
15			8.38							0
76		29.20								0
77		39.66								0

TABLE 8 (Continued)

Off-Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
78			10.76							0
79			10.77							0
84	0									0
85		0								0
86		8.67								0
87	0									0
1				7.75						1.31
2			12.99							0
3			7.02							0
4				15.26						0
5		2.64						0		0
6		15.47						0		0
8		11.89								0
88		11.95								0
10	0									0
12	0									0

TABLE 9

AVERAGE DELAY TIME PER SECTION (SECS)

Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
7		23.63						12.77		0
9				35.60						0.27
36		7.84								0
57		36.88								0
58		10.03								2.16
59	0							0		0
65	0									0
13				16.14				0		6.22
16				17.00				0		0.76
18		8.41								0
73			17.67					0		0
89		13.94						0		1.53
90			20.54					0		0
19			9.39					0		0
20			14.19					0		0
21	0							0		0.49
22	0							0		0

TABLE 9 (Continued)

Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
23			21.93					0		1.07
24			28.00					0		6.88
25		5.96								0
27			10.69						25.60	0
28				23.68					0	0
31			8.65							1.69
32			52.92							0.91
33		8.57						16.70		0
34	0							31.43		2.06
45	0									0
46	0									0
48			12.18							0
49		12.92								0
50		12.36								0
38				17.30						0
40			16.64					26.48		0
41		8.43								0
42		4.45								0
62		0								0

TABLE 9 (Continued)

Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
63			5.32							0
51		11.38								0
52		4.01								0
53		13.05								3.64
54		16.06								0
60	0									0
67	0									0
68	0									0
69	0									0
71	0									0
82		5.88								0
15			13.92							2.63
77		83.29								0
79			24.79							8.56
84	0									0
85		3.56								0
86		21.77								0
87	0									0
1				26.31						0

TABLE 9 (Continued)

Peak

Section Number	Traffic Signals 0	1	2	3	4	5	6	Railroad Crossings 1	2	Misc. Delays Per Mile
2			10.57							0.88
4				14.87						0
5		10.57						0		0
8		17.09								0
88		10.89								2.40
10	0									0
12	0									0

TABLE 10

AVERAGE DELAYS DUE TO TRAFFIC SIGNALS, RAILROAD CROSSINGS, AND OTHER FACTORS (SECS)

Off-Peak Condition

	Traffic Signals							Railroad Crossings		Misc.
	0	1	2	3	4	5	6	1	2	Per Mile
Average Delay	0	9.13	11.91	14.12		12.63	20.99	3.98	7.73	0.40
Number of Sections	22	34	21	11		1	1	24	6	90
Standard Deviation	0	8.43	5.05	9.55		0	0	8.20	13.28	1.25

Peak Condition

	Traffic Signals							RR Xings		Misc.
	0	1	2	3	4	5	6	1	2	Per Mile
Average Delay	0	14.62	17.83	21.56				5.14	25.60	0.68
Number of Sections	16	24	15	7				17	1	62
Standard Deviation	0	16.19	11.60	7.50				10.25	0	1.69

The delay encountered at traffic signals was found to increase as the number of signals increased and as the condition changed from off-peak to peak. Due to the small number of observations involving sections containing four or more signals analysis was limited to those sections containing one, two, and three signals. During the off-peak period the average delay caused by a single signal was about nine seconds while the delay of a section which contained three signals was found to be approximately fourteen seconds. For the peak condition the respective values were approximately fourteen and one-half seconds and twenty-one and one-half seconds. If each condition is examined individually it is found that as the number of traffic signals per section increased the individual effect of each decreased. During the off-peak period the individual effect for one, two, and three signals is 9.13, 5.95, and 4.71 seconds, respectively. Similarly, for the peak period the delays are 14.62, 8.93, and 7.18 seconds, respectively. Intuitively this reflects the effect of those sections which exhibited good progression. The standard deviations that were calculated tend to emphasize the effect of the few signals where the delay for each observation was large because the intersection capacities were exceeded. It was at these intersections that the driver waited through as many as three and four total cycles, whereas at the more typical traffic signal the

wait was only through a portion of a complete cycle. This situation clearly indicates that whenever one is dealing with the condition where the intersection is operating at above capacity, a stop delay study at the particular location will be necessary to obtain accurate estimates of stop delay time.

The average delays encountered in sections which contained one or two railroad crossings (not tracks) were also calculated (Table 10). The standard deviations indicate a high degree of variability in the stop delay. Though a railroad crossing may exist, the number of times that one is delayed is often very few. This is dependent on the frequency that railroad stock uses the particular crossing. Furthermore, if a crossing is blocked, the stop delay of the motorist is dependent on the length and speed of the train. In some cases individual delays at railroad crossings in this study were as long as four or five minutes. To generalize about the magnitude of stop delays at railroad crossings is virtually impossible. It is realistic to conduct individual delay studies at the particular crossings in question to determine these data.

The delays due to the Miscellaneous Delays factor for the off-peak and peak periods were both very small, 0.40 and 0.68 seconds per mile, respectively.

The application of estimated average stop delay values will be most useful in determining and presenting benefits

of traffic engineering improvements. The average values of delay found in this study when utilized according to the restrictions noted enables an estimate of the benefits gained by particular improvements, e.g. constructing a parallel route with fewer traffic signals, the prohibition of left turns or the utilization of one street over another. These benefits could be measured in terms of either time or its chosen money equivalent. The findings also emphasize the fact that the number of traffic signals should be kept to a minimum and that good progression is essential to minimizing stop delay time. A third application of these values might be to point out those sections or intersections which are operating above capacity by contrasting actual delays with the average values found in this study.

CONCLUSION

The primary objective of this project was to formulate a model for predicting the average running speed on urban arterial streets. This has been fulfilled within specified limits of city size and area within the city. From this project comes knowledge of the interaction between frictional factors influencing the motorist and of those frictional factors that most affect him.

The model in its final form is:

$$\hat{Y} = 19.625 + 10.380X_A + 0.076X_B - 0.701X_C - 0.222X_D + 1.258X_F - 0.031X_I + 7.389X_J - 0.186\frac{X_D X_G}{X_A} - 0.242\frac{X_E X_H}{X_A} + 0.018X_I X_K$$

where

X_A = Section length in miles

X_B = Approach width in feet

X_C = Number of signalized intersections

X_D = Number of unsignalized intersections

X_E = Number of alleys

X_F = Population class (30,000 to 50,000 = 0; 50,000 to 100,000 = 1; 100,000 to 150,000 = 2)

X_G = Number of railroad crossings

X_H = Whether or not parking is allowed on the right side of section (No Parking = 0; Parking = 1)

X_I = The fifteen minute volume in number of vehicles during the time the speed is desired

X_J = Lowest G/C ratio of all the signals included in the section

X_K = Whether or not good progression of signals exists (No = 0; Yes = 1).

Refer to Table 4 for exact definition of each variable and for the limitations.

Although the coefficient of multiple determination (R^2) was not as high as might be desired and although difficulties were encountered in the testing of some of the assumptions, it is believed that the model as a predictor has merit. Estimates of average running speed within a few miles per hour of the true value will be sufficient for most planning and traffic engineering uses. The proposed model has a standard error of estimate of 3.417 mph. For ten examples the model estimated the average running speed within approximately four miles per hour and estimated fifty percent of the values within two miles per hour.

As one reads the list of variables utilized in the model it becomes noticeable that some of these input variables are identical to ones used in capacity calculations. This is what was desired at the outset of the project.

It is of interest to evaluate the individual effects on average running speed of the variables. Section length affects this speed more than any other variable. The difference in running speed between two sections whose only difference is a half mile in length is about five miles per hour. A variable which substantially affects running speed is volume. The speed is reduced 1.3 mph for every one hundred vehicle increase. Moreover, if a section contains more than one signalized intersection with poor progression, the reduction increases to 3.1 mph for the same number of vehicles.

The population of the city increases the speed to a greater extent than anticipated. For every fifty thousand people increase the speed correspondingly increases 1.3 mph, emphasizing the fact that people drive faster on this type of facility in larger cities.

An important fact contained in the model is that small differences in the lowest G/C ratio produce substantial differences in the speed. As an example, a section whose lowest G/C ratio is 0.50 would have a predicted speed seven-tenths of a mile per hour slower than a section whose lowest G/C ratio is 0.60. The approach width of the section also substantially affects the average running speed. The difference in speed between a section having a twelve foot approach width and a second section having a twenty-four foot approach width is nine-tenths of a mile per hour.

The addition of a traffic signal within a section decreases speed by about one-half mile per hour over that existing before installation of the traffic signal. A further reduction results if there are two or more signals in the section and the progression is unsatisfactory. Table 11 is a summary of this paragraph.

To supplement the usefulness of the proposed prediction model, the stop delay data collected was also analyzed. This research project found the major cause of stop delay to be the traffic signal. In this study the second most important cause of stop delay was due to railroad crossings. These delays increased as the number per section increased for both the traffic conditions of off-peak and peak. The percentage of the total stop delay caused by traffic signals was eighty to eighty-five percent while the percentage caused by railroad crossings was twelve to fifteen percent. Stop delay due to other causes was less than three percent.

Considering sections which contained one traffic signal, the average stop delay was 9.13 seconds during the off-peak period and 14.62 seconds for the peak period. Other sections containing three traffic signals had average stop delays of 14.12 seconds and 21.56 seconds, respectively.

TABLE 11

SUMMARY OF SPEED INFLUENCING FACTORS

Variable Name	Variable Increase	Speed Change
Section Length	0.1 mile	+ 1.0 mph
Approach Width	12 feet	+ 0.9 mph
Signalized Inter.	1 unit	- 0.7 mph
Unsignalized Inter.	1 unit	- 0.2 mph
Population	50000 people	+ 1.3 mph
Volume	100 vehicles	- 3.1 mph (poor progression)
		- 1.3 mph (good progression)
Lowest G/C Ratio	10 percent	+ 0.7 mph

Using both the prediction model for average running speed and the average values of stop delay enables one to determine the location and extent of improvements to the travel time and to estimate the benefits gained from them. Comparing the benefits with the costs could indicate whether or not the proposed improvement was worthwhile. Alone, the prediction model would function as a means of determining which frictional factors are reducing speed and to what degree so that corrections might be made. Another function of the model would be the determination of appropriate speed limits on urban arterial streets. Hopefully, the knowledge contained in this report will be put to use in the field of traffic engineering.

BIBLIOGRAPHY

1. Campbell, E. W., Keefer, L. E., and Adams, R. W., "A Method for Predicting Speeds Through Signalized Sections." H.R.B. Bulletin 230, (1959).

2. Draper, N. R. and Smith, H., Applied Regression Analysis. New York: John Wiley & Sons, Inc., 1966.

3. Edwards and Kelcey, Inc., "Optimizing Flow on Existing Street Networks." Highway Research Board, N.C.H.R.P. Project 3-14, (1969).

4. Foster, L. A., Testing for Equality of Variances. Diss. Purdue 1964.

5. Greenshields, B. D., "Measurement of Highway Traffic Performance." Traffic Engineering, XXXIX (April 1969), 26-30.

6. Greenshields, B. D. and Weida, F. M., Statistics With Applications to Highway Traffic Analysis. Saugatuck: Eno Foundation for Highway Traffic Control, 1952.

7. Hejal, S. S. and Michael, H. L., "A Predictive Model for Travel Times on Two-Lane Highways." J.H.R.P. No. 19, (1970).

8. Helly, W. and Baker, P. G., "Acceleration Noise in a Congested Signalized Environment." Vehicular Traffic Science, (1967).

9. Highway Research Board, "Highway Capacity Manual." H.R.B. Special Report 87, (1965).

10. Institute of Traffic Engineers, Manual of Traffic Engineering Studies. Washington: Institute of Traffic Engineers, 1964.

11. Jones, T. R. and Potts, R. B., "The Measurement of Acceleration Noise - A Traffic Parameter." Operations Research, X (1962), 745-763.

12. Lilliefors, H. W., "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown." Journal of the American Statistical Association, LXII (June 1967), 400.

13. Matson, T. M., Smith, W. S., and Hurd, F. W., Traffic Engineering. New York: McGraw-Hill, 1955.

14. Miller, I. and Freund, J. E., Probability and Statistics for Engineers. Englewood Cliffs: Prentice-Hall, Inc., 1965.

15. Ostle, B., Statistics in Research. Ames: The Iowa State University Press, 1963.

16. Snedecor, G. W. and Cochran, W. G., Statistical Methods. Ames: The Iowa State University Press, 1967.

17. Statistical Laboratory Library Programs, Purdue University.

18. Votaw, D. F. and Levinson, H. S., Elementary Sampling for Traffic Engineers. Saugatuck: Eno Foundation for Highway Traffic Control, 1962.

19. Webster, L. A., "Traffic Delay on Urban Arterial Streets as a Result of Curb Parking Maneuvers." H.R.B. Record 267, (1967).

20. Wohl, M. and Martin, B. V., Traffic System Analysis for Engineers and Planners. New York: McGraw-Hill, 1967.

APPENDIX A

DISCUSSION OF THE FOSTER-BURR Q-TEST

THE Q-TEST FOR EQUALITY OF VARIANCES*

by

Louis A. Foster and Irving W. Burr

The Q-test for equality of variances is based on a statistic which is a monotone function of the coefficient of variation of the sample variances.** As such it offers promise as a preliminary test for the assumption of homogeneity of population variances which is needed in the analysis of variance technique.*** Although the Q-test is not a so-called quick test, the test statistic is sufficiently simple to permit calculation of its value on a desk calculator (in contrast to the logarithmic transformation required in Bartlett's test). A sample variance taking the value zero does not disrupt this test (as it does to Bartlett's and Hartley's tests). Finally, a table of critical values permits the use of this test for small sample sizes where the use of an asymptotic distribution would not be appropriate.

THE Q-STATISTIC: For equal sample sizes, n, from each of p parent populations, let s_j^2 (for $j = 1,\ldots,p$) denote the jth sample variance. Denoting the value of the test statistic Q by q, we have:

$$q = (s_1^4 + \ldots + s_p^4)/(s_1^2 + \ldots + s_p^2)^2.$$

For unequal sample sizes we specify that each sample variance, s_j^2, be calculated by dividing by the degrees of freedom, v_j, rather than by the sample size, n_j (where

---

*This brief summary of the Q-test is based on a Ph.D. thesis submitted to Purdue University by Louis A. Foster working under the supervision of Irving W. Burr.

**Equivalent tests have been proposed by Brandt and Stevens, but in a form depending solely on an asymptotic distribution for the test statistics.

***This statement is based on the role of the coefficient of variation of the population variances in measuring the disruption of the F-ratio in the analysis of variance technique, as set forth by Box (6) Vol. 25, pages 290 and 484.

$v_j = n_j - 1$, for $j = 1,\ldots,p$). Let $\bar{v}$ denote the arithmetic average of the degrees of freedom. In this case we have:

$$q = \bar{v}\ (v_1 s_1^4 + \ldots + v_p s_p^4)/(v_1 s_1^2 + \ldots + v_p s_p^2)^2.$$

THE Q-TEST: Large values of Q lead to rejection of the hypothesis of equal population variances. The critical values are given in the attached table for various numbers of parent populations, p, and various possibilities for equal degrees of freedom, v (where $v = n - 1$), from one to ten. This table can be used directly for equal degrees of freedom. For unequal degrees of freedom, q is calculated as indicated above, but $\bar{v}$ is to be substituted for v in the attached table, provided that $\bar{v}$ and the harmonic mean of the v's do not differ greatly. For large sample sizes ($v > 10$), we note that $p\ \bar{v}\ (pQ-1)/2$ is asymptotically chi-square with (p-1) degrees of freedom.

The 95th and 99th Percentiles of Q for Equal Degrees of Freedom and Equal Population Variances

(the upper entry in each cell is to be used for .05 level tests and the lower entry for .01 level tests).

The Degrees of Freedom (v = n-1)

	1	2	3	4	5	6	7	8	9	10
2	.999	.951	.886	.829	.785	.750	.722	.700	.681	.666
	.9999	.990	.960	.921	.822	.848	.818	.792	.770	.751
3	.923	.747	.658	.594	.548	.520	.497	.479	.464	.453
	.999	.857	.781	.703	.649	.607	.575	.550	.529	.512
4	.814	.612	.517	.485	.422	.396	.377	.362	.350	.341
	.994	.744	.628	.549	.498	.461	.434	.413	.396	.383
5	.708	.510	.421	.371	.339	.318	.302	.290	.280	.272
	.856	.630	.513	.443	.399	.368	.346	.328	.314	.304
6	.662	.433	.353	.310	.283	.264	.251	.241	.233	.226
	.770	.540	.430	.369	.331	.305	.285	.271	.260	.250
8	.495	.331	.266	.232	.211	.297	.187	.179	.173	.168
	.634	.413	.322	.274	.244	.224	.210	.199	.191	.184
9	.449	.295	.236	.206	.187	.175	.166	.159	.153	.149
	.577	.367	.285	.241	.215	.198	.185	.176	.168	.162
10	.410	.266	.212	.284	.168	.156	.148	.142	.138	.134
	.527	.330	.257	.216	.192	.176	.165	.157	.150	.145
12	.348	.221	.176	.153	.139	.129	.123	.118	.114	.111
	.448	.273	.209	.177	.156	.145	.136	.129	.124	.119
16	.264	.164	.130	.113	.103	.0958	.0910	.0874	.0845	.0823
	.339	.200	.152	.129	.115	.106	.0991	.0944	.0907	.0877

	1	2	3	4	5	6	7	8	9	10
27	.155	.0944	.0747	.0650	.0592	.0554	.0527	.0507	.0492	.0479
	.195	.112	.0849	.0721	.0646	.0597	.0563	.0538	.0518	.0503
32	.129	.0787	.0623	.0543	.0495	.0464	.0442	.0425	.0413	.0402
	.161	.0920	.0702	.0597	.0537	.0497	.0469	.0449	.0433	.0420
64	.0613	.0375	.0299	.0262	.0240	.0226	.0216	.0208	.0202	.0197
	.0731	.0420	.0325	.0280	.0254	.0236	.0224	.0216	.0209	.0203

APPENDIX B

NORMALITY TEST FOR SECTION SPEEDS

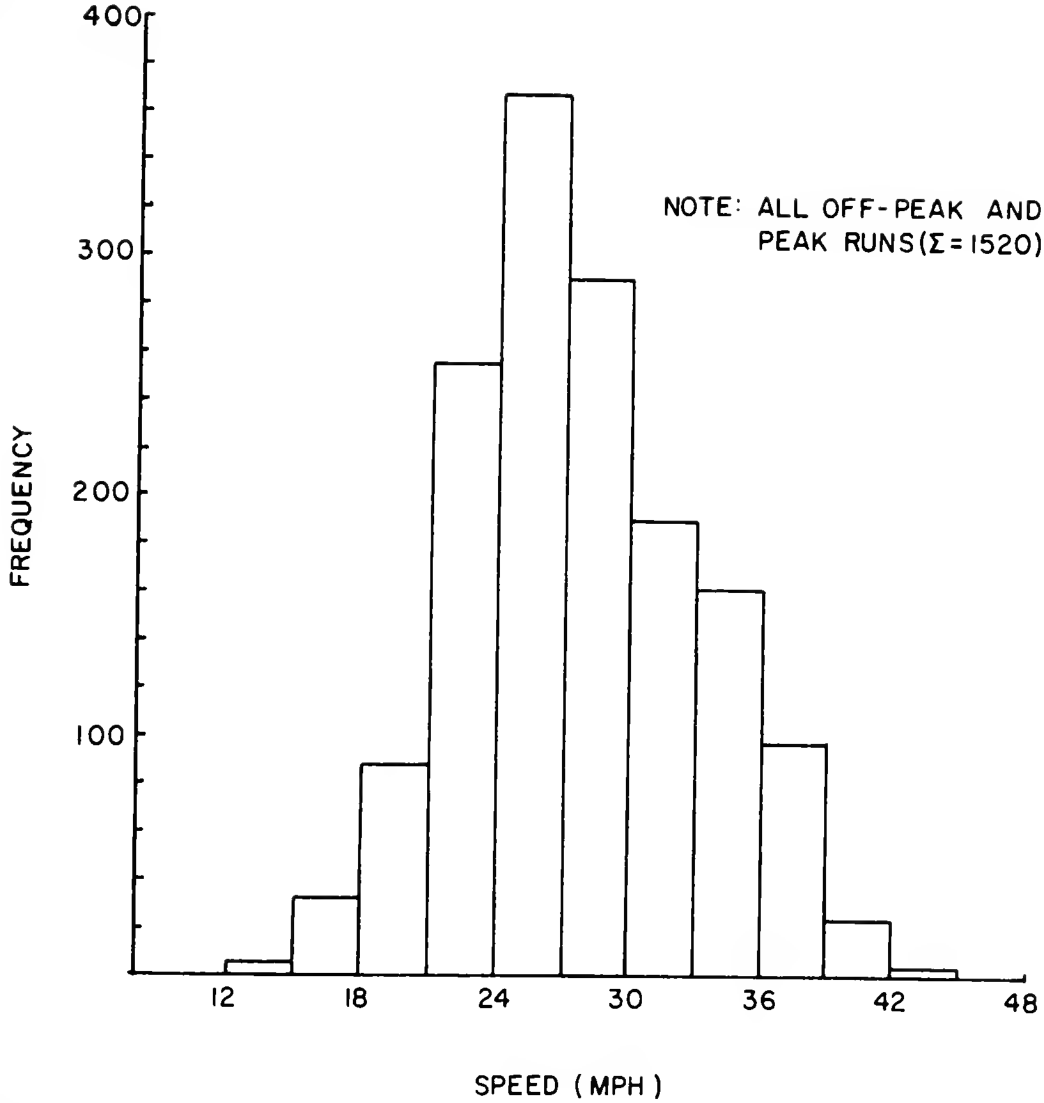

FIGURE B1. FREQUENCY OF SECTION SPEEDS

TABLE B1

NORMALITY TEST FOR SECTION SPEEDS (OFF-PEAK AND PEAK)

Class	Frequency	Cumulative Frequency	Cumulative Frequency (%)	Z	Normal Curve Cumulative Frequency (%)
12.000-14.999	6	6	0.4	-2.39	0.842
15.000-17.999	32	38	2.5	-1.83	3.362
18.000-20.999	88	126	8.3	-1.26	10.383
21.000-23.999	255	381	25.1	-0.70	24.196
24.000-26.999	368	749	49.3	-0.14	44.433*
27.000-29.999	291	1040	68.4	+0.42	66.276
30.000-32.999	191	1231	81.0	+0.99	83.891
33.000-35.999	162	1393	91.6	+1.55	93.943
36.000-38.999	98	1491	98.1	+2.11	98.257
39.000-41.999	25	1516	99.7	+2.67	99.621
42.000-44.999	4	1520	100.0	+3.23	99.938

*Denoted Location of Maximum Deviation.

Calculations for the Kolmogorov-Smirnov Test
for the Normality of the Section Speeds

Critical D for 1520 samples (Level of Significance = 0.01)

$$D_c = \frac{1.031}{(N)^{1/2}} \quad (12)$$

$$D_c = \frac{1.031}{(1520)^{1/2}} = 0.0264$$

Maximum Deviation = 0.0487

0.0487 > 0.0264

Therefore, reject hypothesis that the curve is normal.

APPENDIX C

NORMALITY TEST FOR RESIDUALS

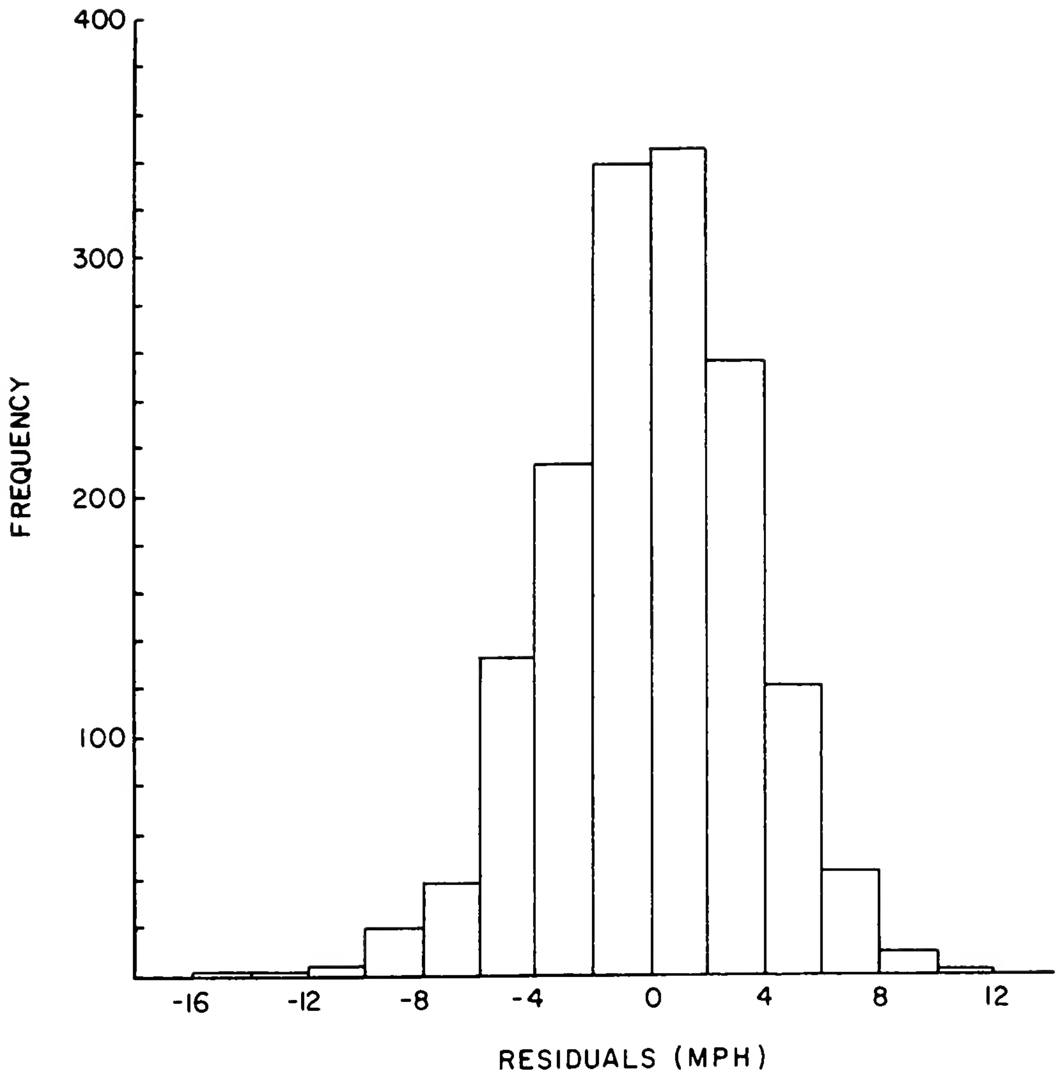

FIGURE C1. FREQUENCY OF RESIDUALS

TABLE C1

NORMALITY TEST FOR RESIDUALS

Class	Frequency	Cumulative Frequency	Cumulative Frequency (%)	Z	Normal Curve Cumulative Frequency (%)
-16 to -14	1	1	0.1	-4.10	--
-14 to -12	1	2	0.1	-3.52	0.022
-12 to -10	3	5	0.3	-2.93	0.169
-10 to -8	19	24	1.6	-2.34	0.964
-8 to -6	38	62	4.1	-1.76	3.920
-6 to -4	132	194	12.8	-1.17	12.100
-4 to -2	213	407	26.8	-0.59	27.760*
-2 to 0	339	746	49.1	0	50.000
0 to 2	345	1091	71.8	0.59	72.240
2 to 4	256	1347	88.5	1.17	87.900
4 to 6	120	1467	96.5	1.76	96.080
6 to 8	43	1510	99.3	2.34	99.036
8 to 10	9	1519	99.9	2.93	99.831
10 to 12	1	1520	100.0	3.52	99.978

*Denotes location of maximum deviation.

Claculations for the Kolmogorov-Smirnov Test
for the Normality of the Residuals

Critical D for 1520 samples (Level of Significance = 0.05)

$$D_c = \frac{0.886}{(N)^{1/2}} \qquad (12)$$

$$D_c = \frac{0.886}{(1520)^{1/2}} = 0.0228$$

Maximum Deviation = 0.0096

$0.0096 < 0.0228$

Therefore, cannot reject hypothesis that the curve is normal.